* ***Agricultural intensity map***

**Global Land Cover and Land Use 2000 and 2020 (30 m resolution)**

dataset:<https://storage.googleapis.com/earthenginepartners-hansen/GLCLU2000-2020/v2/download.html>

paper:<https://www.frontiersin.org/journals/remote-sensing/articles/10.3389/frsen.2022.856903/full>

[2000](https://storage.googleapis.com/earthenginepartners-hansen/GLCLU2000-2020/v2/2000.txt), [2005](https://storage.googleapis.com/earthenginepartners-hansen/GLCLU2000-2020/v2/2005.txt), [2010](https://storage.googleapis.com/earthenginepartners-hansen/GLCLU2000-2020/v2/2010.txt), [2015](https://storage.googleapis.com/earthenginepartners-hansen/GLCLU2000-2020/v2/2015.txt), [2020](https://storage.googleapis.com/earthenginepartners-hansen/GLCLU2000-2020/v2/2020.txt). [2000-2020change](https://storage.googleapis.com/earthenginepartners-hansen/GLCLU2000-2020/v2/2000-2020change.txt).

All 2020 tiles have been projected to Mollweide, under: “*sparing\_sharing\sparing\_rproject\Data\GLCLU\_2020\_projected”*

The legend for this dataset is under:

*“sparing\_sharing\sparing\_rproject\Data\GLCLU legend.xlsx”*

(Class 244 in orange is cropland.)

***Cultivated land data (30m)***

dataset:<https://stac.openlandmap.org/gpw_ggc-30m/gpw_ggc-30m_20200101_20201231/gpw_ggc-30m_20200101_20201231.json?.asset=asset-gpw_cultiv.grassland_rf.med.filt_p_30m>

paper: <https://www.nature.com/articles/s41597-024-04139-6#Sec15>

This layer has been projected to Mollweide as well, under: *"sparing\_sharing\sparing\_rproject\Data\mo\_cultivated\_2020.tif"*

**Agricultural = cropland from GLCLU + cultivated land (exclude built-up and water)**

Original agricultural intensity at 30m tiles are under:

*"sparing\_sharing\sparing\_rproject\Data\glclu\_with\_cultivated\_2020”*

Then I aggregate them into 300m, under:

*"sparing\_sharing\sparing\_rproject\Data\with\_cultivated\_intensity\_300m"*

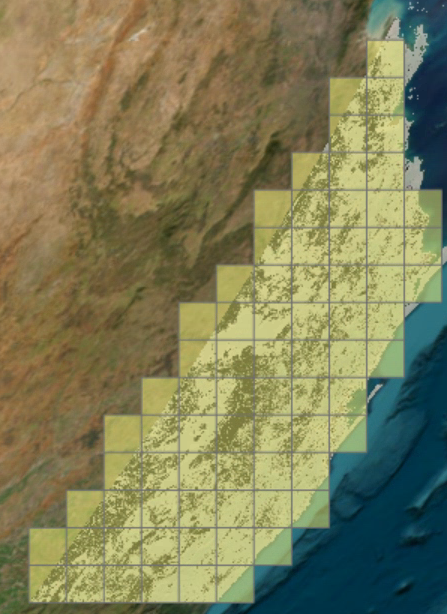
All tiles are in Mollweide projection

* ***QLD tile 600m test***

Based on the 300m agricultural intensity layers, aggregate them into 600m grid size, then see the belta distribution at 60km to classify the 60km grid is sparing or sharing.

Using this function that Jonathan wrote, under: *"sparing\_sharing\sparing\_rproject\classify\_spare\_share.r"*

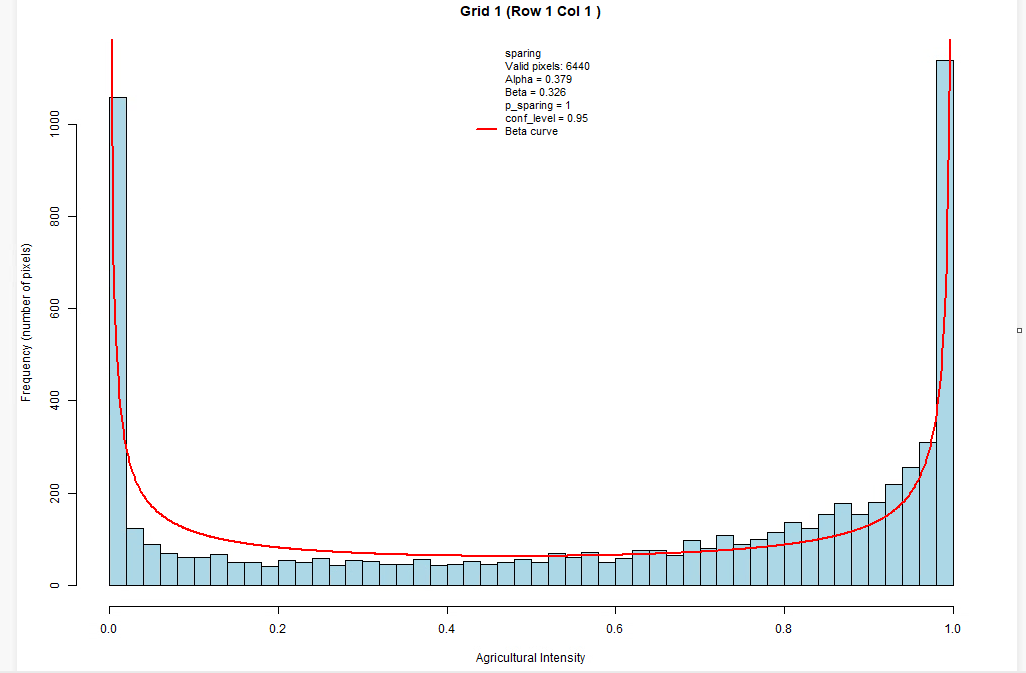
I’ve generated the grid net over the QLD agricultural-intensity tile, I added a filter to remove grids with fewer than 100 valid pixels (1% of 10,000 total pixels in one 60km grid) before running the beta-distribution function, so the ocean/invalid grids are no longer falsely classified (which was the issue in the earlier code):



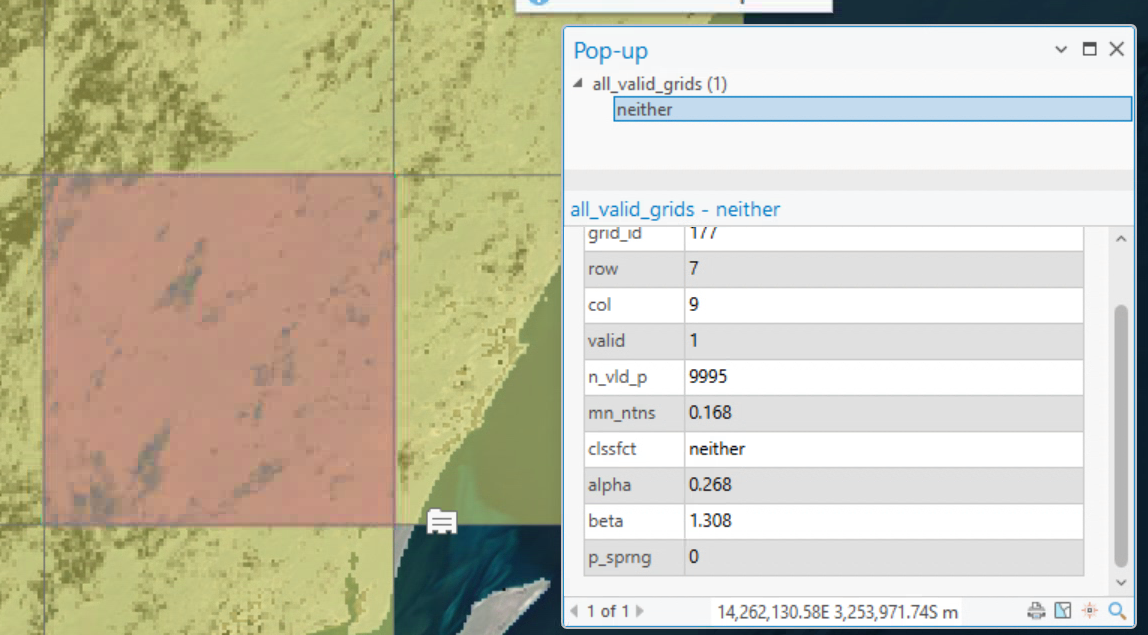
In total 75 valid grids, there is 17 neither, 58 sparing, still no sharing.

For example:

* The grid in the bottom-left corner looks like sparing (black is high agri intensity), and has been correctly identified as **sparing**. Its histogram now shows the correct pixel counts on the y-axis:



* Another grid that looks like it might be **sharing** (highlighted in red) turns out to be classified as **neither**, as α < 1:





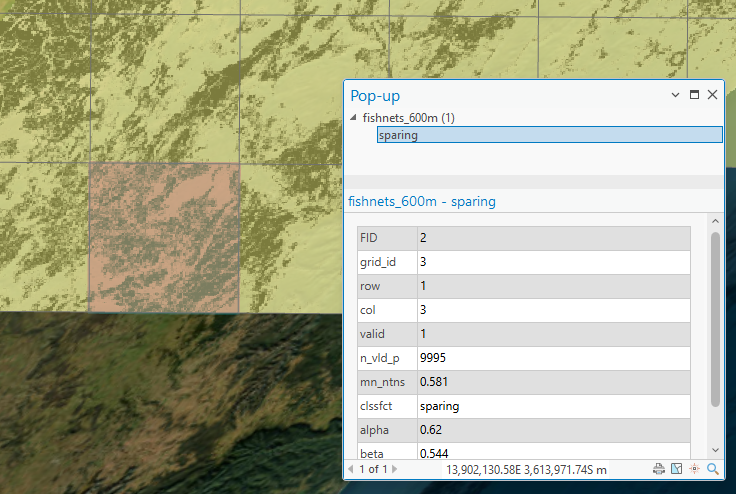
* ***Small grid size scale test***

Based on the 300m agricultural intensity layers, aggregate them into different grid sizes (600m/1200m/2400m), then see the belta distribution at 60km to classify the 60km grid is sparing or sharing.

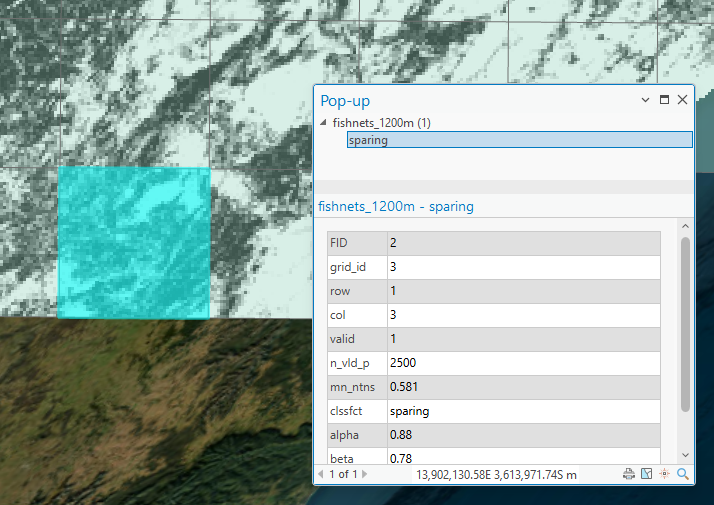
Using the Queensland tile as an example:

* 600 m: 75 valid grids — 58 sparing, 0 sharing, 17 neither
* 1200 m: 75 valid grids — 56 sparing, 0 sharing, 19 neither (very similar to 600 m, with ~75% agreement)
* 2400 m: 74 valid grids — 18 sparing, 6 sharing! 50 neither

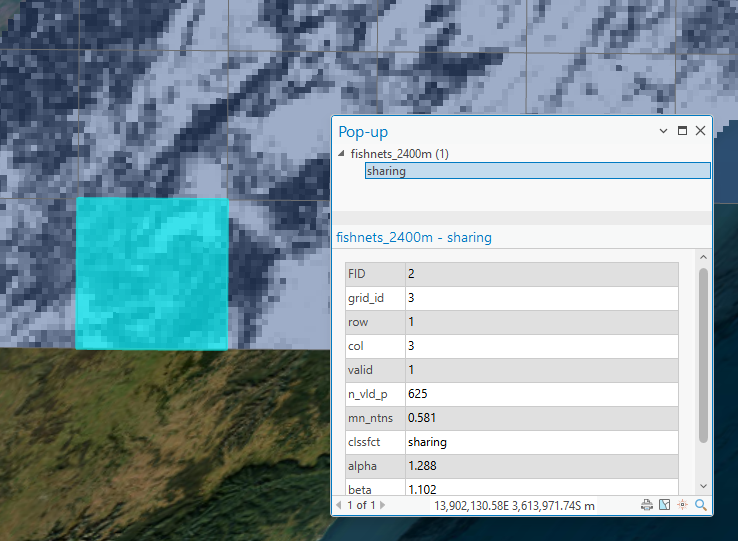
Looking at individual grids (e.g. row 1, column 3), at 600m, it's sparing:

****

at 1200m, it's also sparing:

****

at 2400m, it turns into sharing:

****

All results are stored under: *"sparing\_sharing\sparing\_rproject\beta\_distribution\_method\fishnets\_at\_different\_scale"*

Each grid-size folder contains 1) the fishnet shapefile, 2) histograms of the associated beta distribution results (you can click a grid in ArcGIS to view its row/column ID, then find the corresponding histogram in this folder), and 3) the agricultural intensity map at that grid size.

I’ve also included a summary file

*"Sparing\_sharing\sparing\_rproject\beta\_distribution\_method\fishnets\_at\_different\_scale\resolution\_comparison.csv"*

which shows ~75% agreement and r > 0.99 parameter correlation between 600 m and 1200 m, but only ~46% agreement between 600 m and 2400 m.

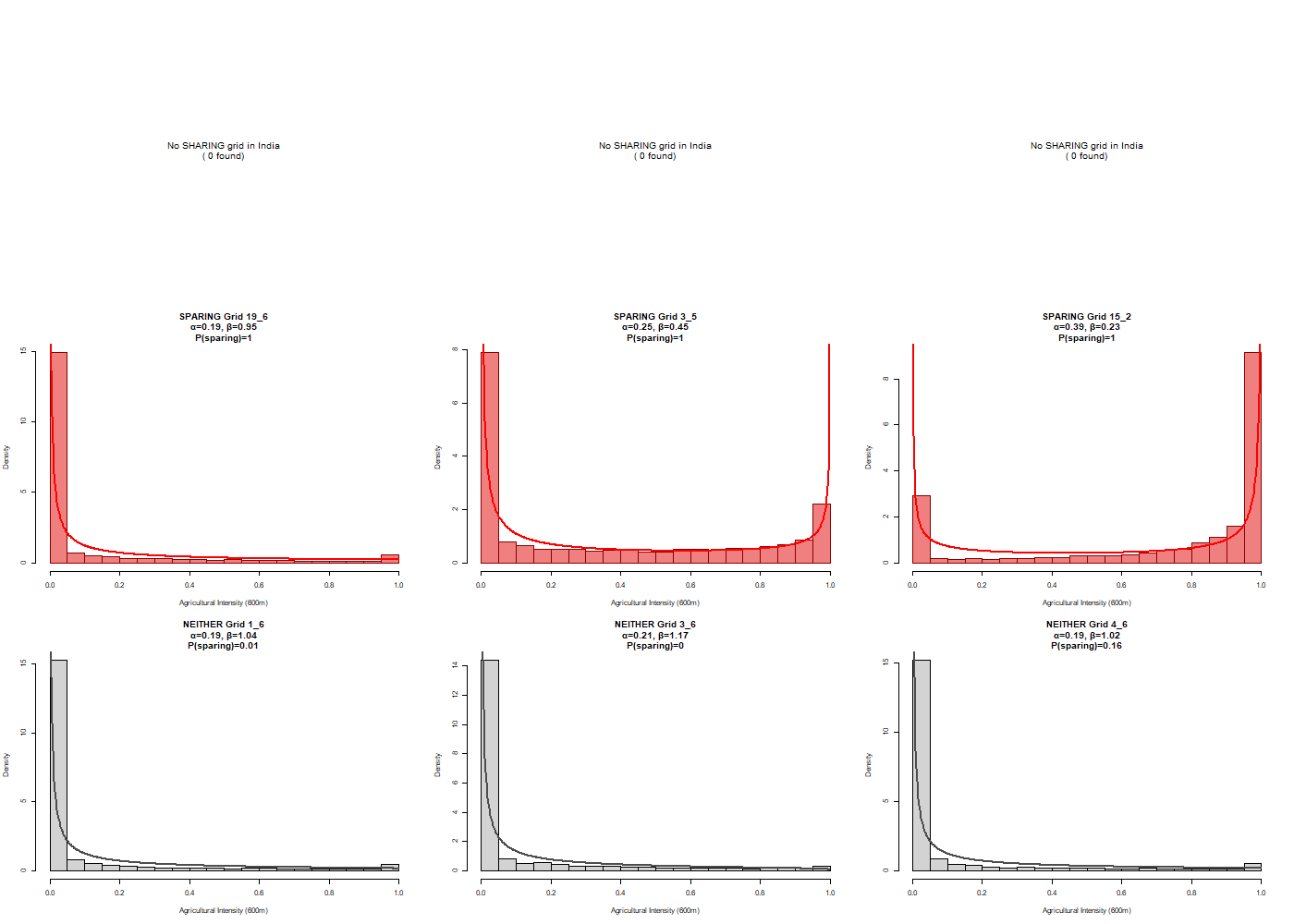
**The rest is from previous abandoned tests, don’t look at it**

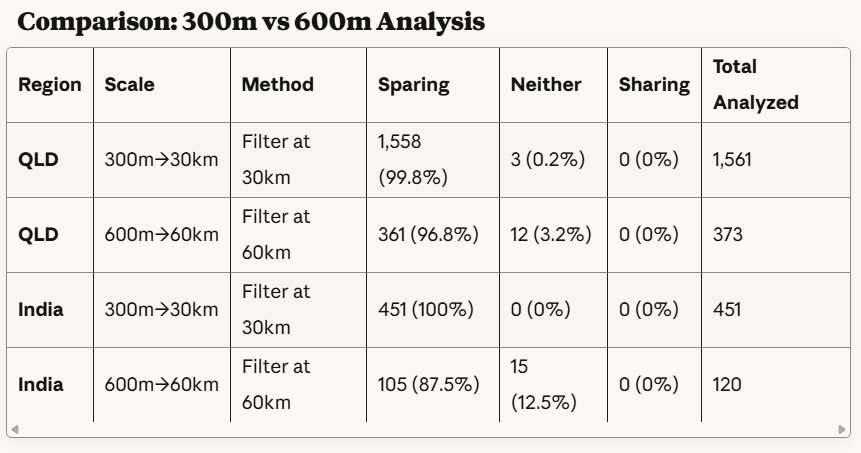
****

* **Adding 10% filter to India tile for belta distribution (filter 60km grid < 0.1 intensity)**

105 sparing, 125 neither, 0 sharing (total: 120 analysed)

Removed 220 grids that had <10% agriculture, 87.5% is sparing

****

****

**"Y:\sparing\_sharing\sparing\_rproject\Data\GLCLU\_2020\_projected\20S\_150E\_m30m.tif"**

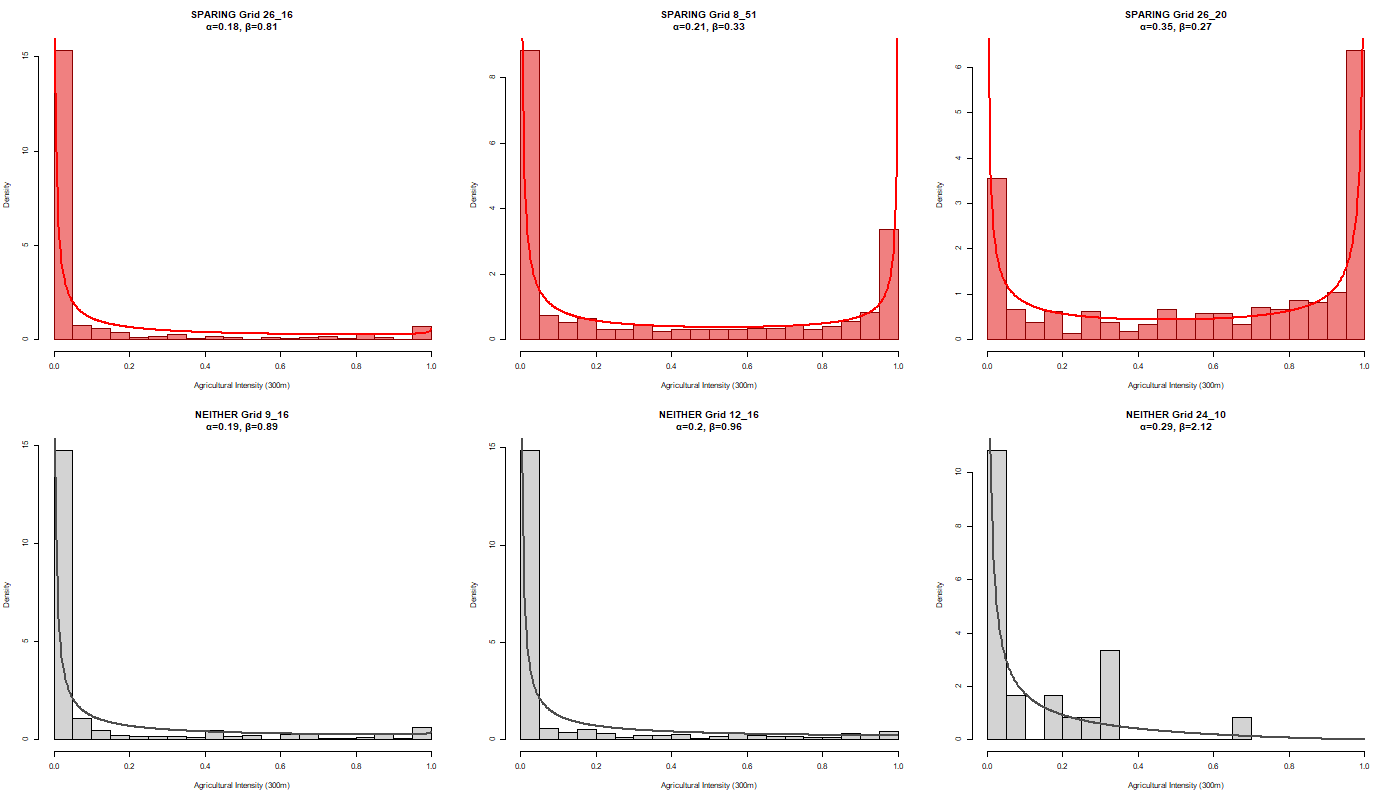
**"Y:\sparing\_sharing\sparing\_rproject\beta\_distribution\_method\scale\_test\qld\_600m\_filter60km.tif"  
  
"Y:\sparing\_sharing\sparing\_rproject\Data\GLCLU legend.xlsx"**

* **Adding 10% filter to QLD tile for belta distribution (filter 30km grid < 0.1 intensity)**

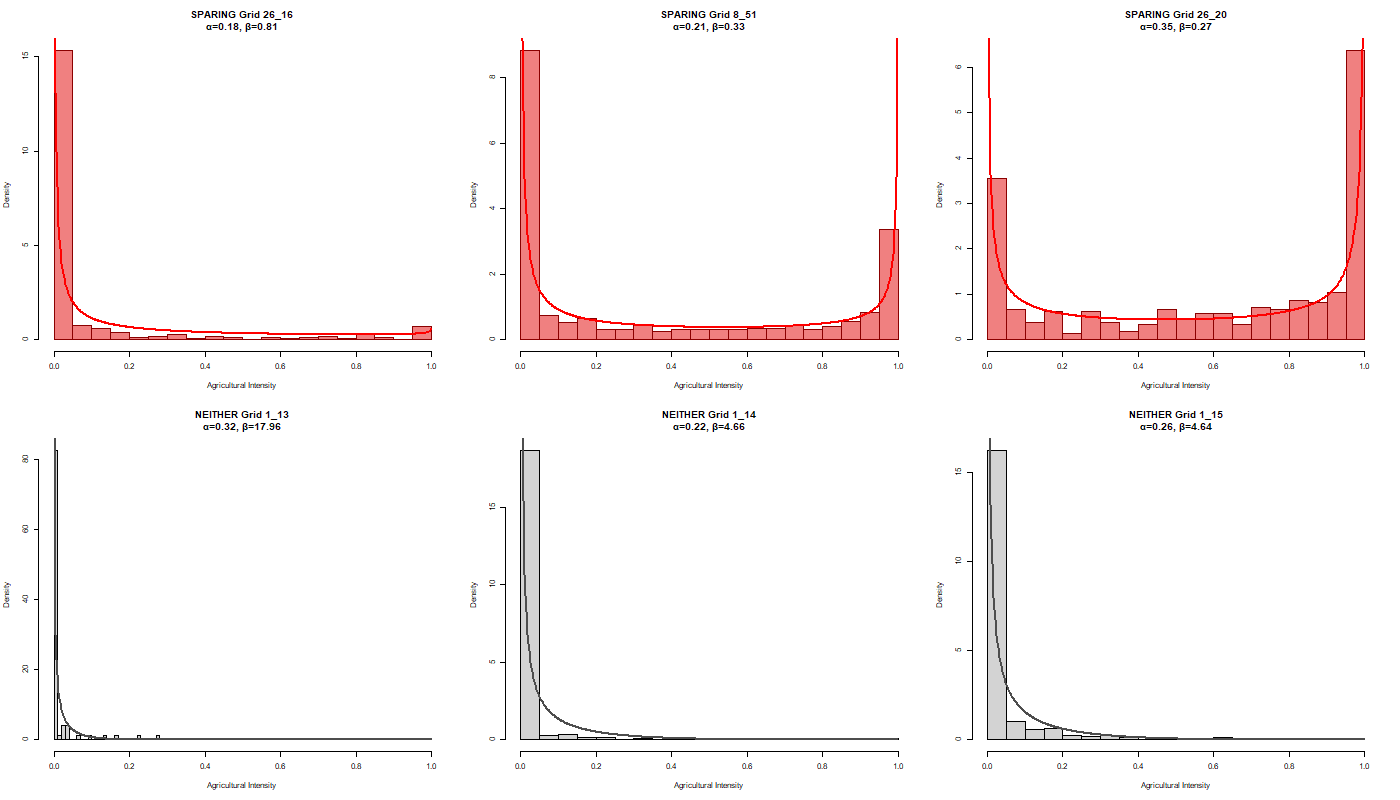
Before filter: 1558 sparing, 130 neither, 0 sharing (total: 1688 analysed)

After filter: 1558 sparing, 3 neither, 0 sharing (total: 1561 analysed)

Removed 127 grids that had <10% agriculture AND removed 127 "neither" grids (130→3). The filter removed low-agriculture "neither" grids, but all remaining grids are STILL sparing (99.8%).



Before filter:

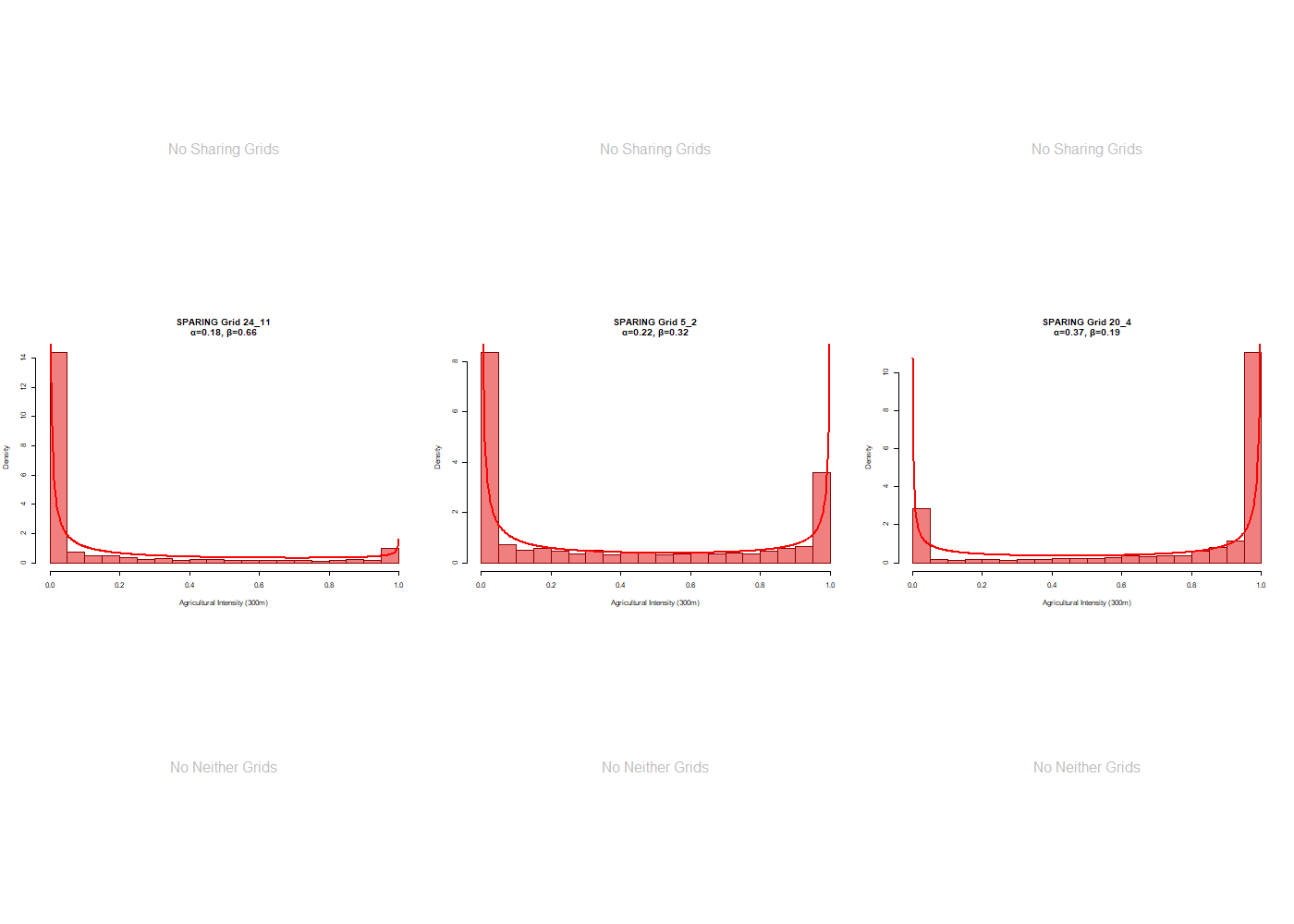


* **India tile (filter 30km grid < 0.1 intensity)**

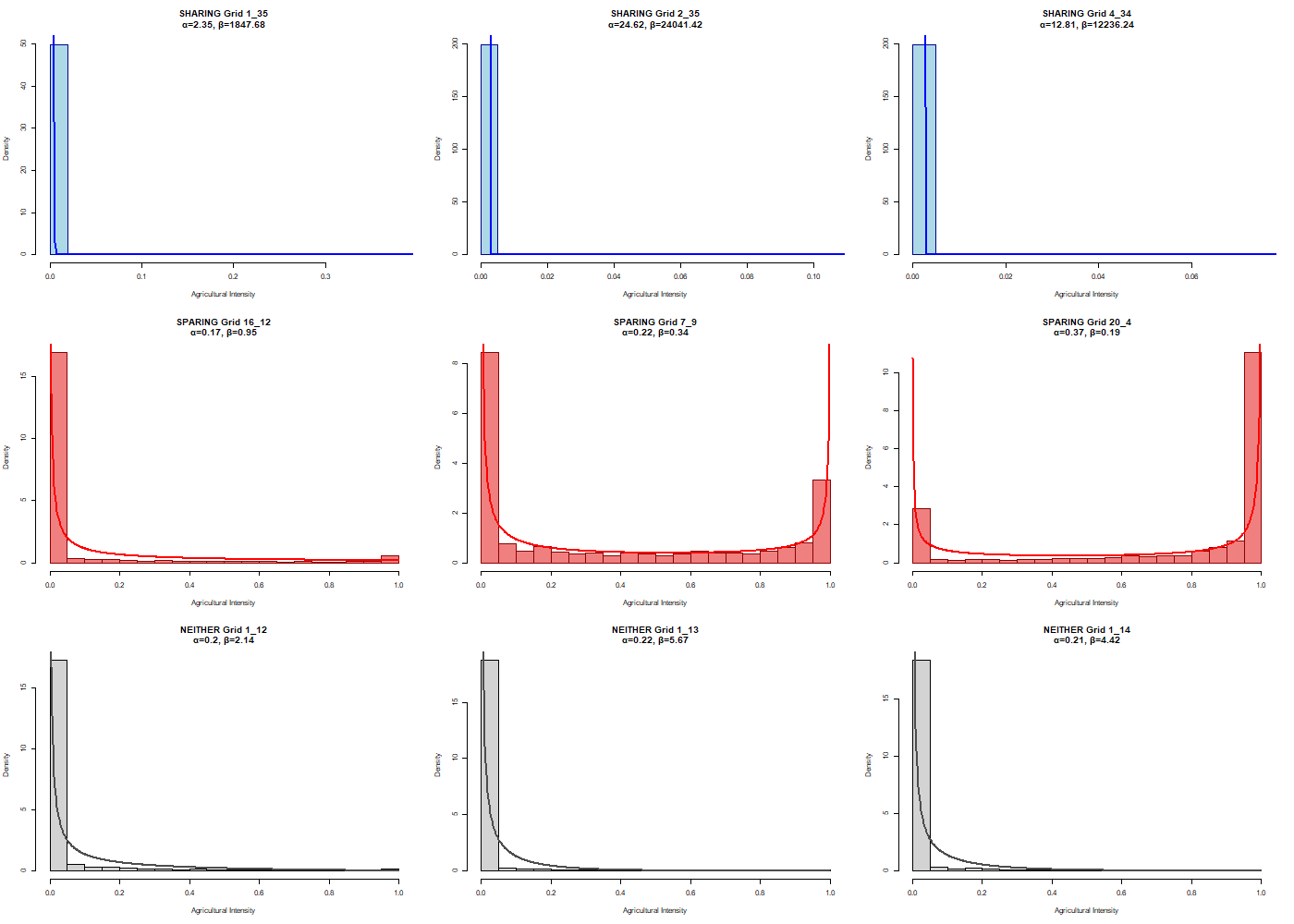
Before filter: 457 sparing, 953 neither, 24 sharing (total: 1434 analyzed)

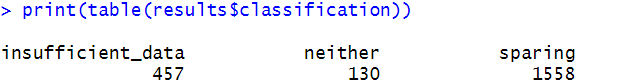
After filter: 451 sparing, 0 neither, 0 sharing (total: 451 analyzed)

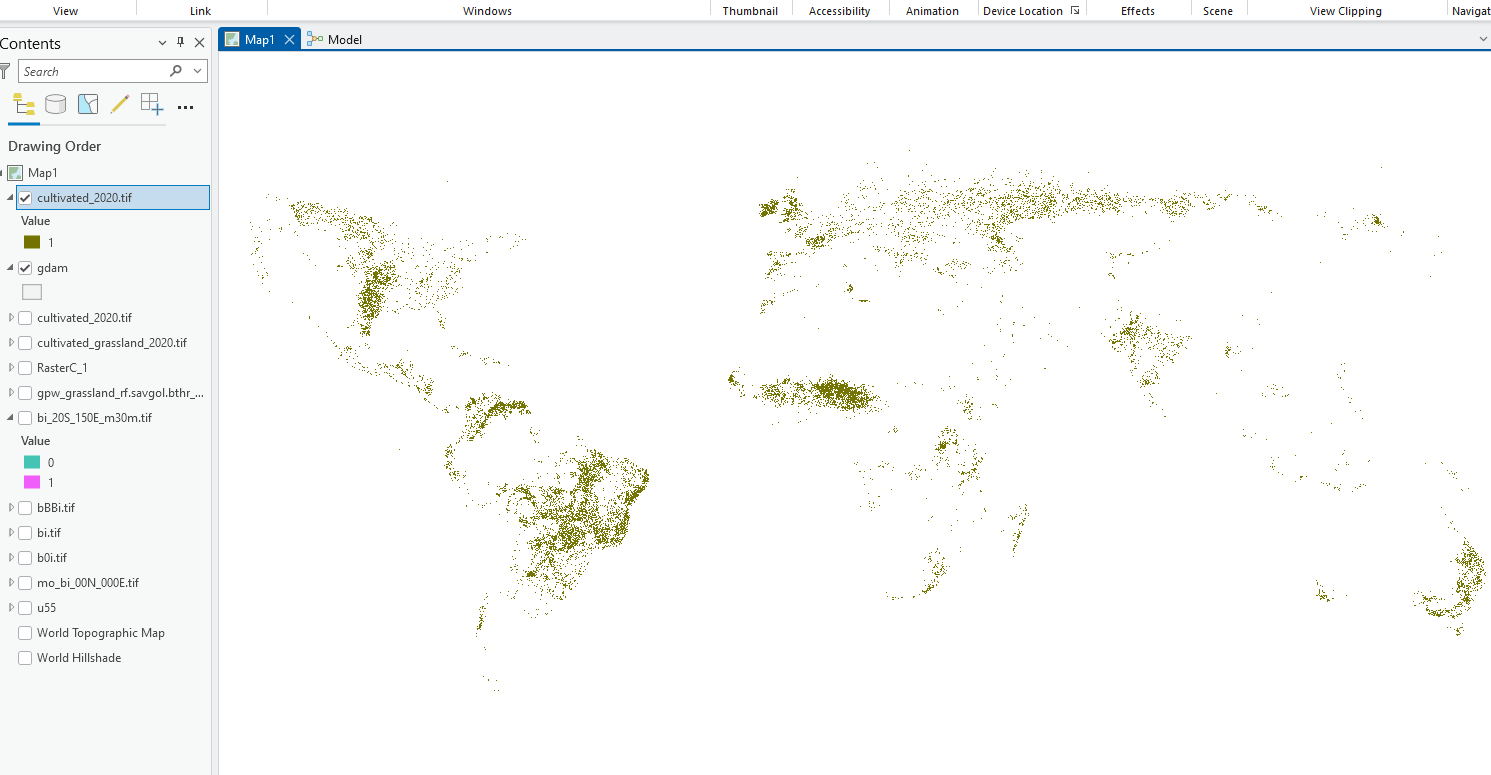
Removed 984 low-agriculture grids including ALL 953 "neither" and all 24 "sharing" grids, no sharing patterns remain.

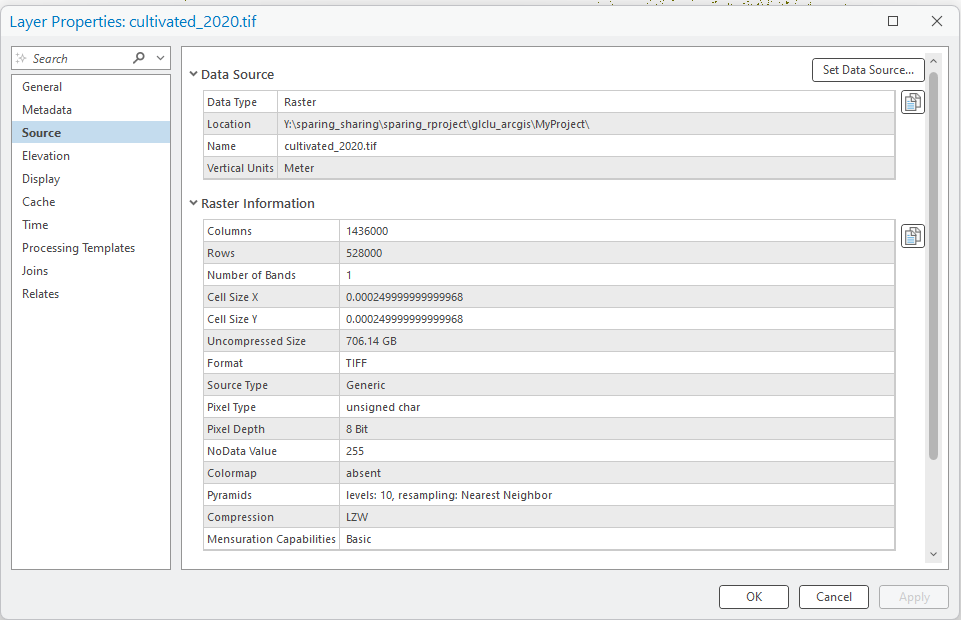
****

**Before filter:**

****

****

****

****

**Global Land Cover and Land Use 2000 and 2020 (30 m resolution)**

dataset:[**https://storage.googleapis.com/earthenginepartners-hansen/GLCLU2000-2020/v2/download.html**](https://storage.googleapis.com/earthenginepartners-hansen/GLCLU2000-2020/v2/download.html)

paper:[**https://www.frontiersin.org/journals/remote-sensing/articles/10.3389/frsen.2022.856903/full**](https://www.frontiersin.org/journals/remote-sensing/articles/10.3389/frsen.2022.856903/full)

[**2000**](https://storage.googleapis.com/earthenginepartners-hansen/GLCLU2000-2020/v2/2000.txt)**,** [**2005**](https://storage.googleapis.com/earthenginepartners-hansen/GLCLU2000-2020/v2/2005.txt)**,** [**2010**](https://storage.googleapis.com/earthenginepartners-hansen/GLCLU2000-2020/v2/2010.txt)**,** [**2015**](https://storage.googleapis.com/earthenginepartners-hansen/GLCLU2000-2020/v2/2015.txt)**,** [**2020**](https://storage.googleapis.com/earthenginepartners-hansen/GLCLU2000-2020/v2/2020.txt)**.** [**2000-2020change**](https://storage.googleapis.com/earthenginepartners-hansen/GLCLU2000-2020/v2/2000-2020change.txt)**.**

**Global cropland expansion- consistent cropland extent time-series at 30-m spatial resolution**

2000-2003, 2004-2007, 2008-2011, 2012-2015, and 2016-2019

dataset:[**https://glad.umd.edu/dataset/croplands**](https://glad.umd.edu/dataset/croplands)

paper:[**https://www.nature.com/articles/s43016-021-00429-z#Sec1**](https://www.nature.com/articles/s43016-021-00429-z#Sec1)

**A global open-source dataset of monthly irrigated and rainfed cropped areas (MIRCA-OS) for the 21st century (10 km resolution)**

**23 crops harvested area, rainfed/irrigated**

dataset:[**https://www.hydroshare.org/resource/60a890eb841c460192c03bb590687145/**](https://www.hydroshare.org/resource/60a890eb841c460192c03bb590687145/)

paper:[**https://www.nature.com/articles/s41597-024-04313-w#Sec1**](https://www.nature.com/articles/s41597-024-04313-w#Sec1)

**Spatial Production Allocation Model (SPAM 2020 V1r0) (10 km resolution)**

website:<https://mapspam.info/>

Methodology paper:<https://www.sciencedirect.com/science/article/pii/S0308521X14000110?via%3Dihub>

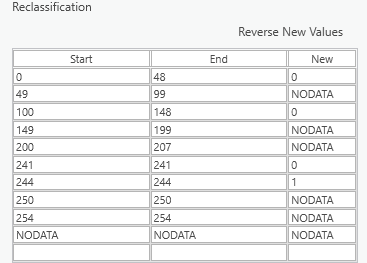
Paper: A cultivated planet in 2010 – Part 2: The global gridded agricultural-production maps:

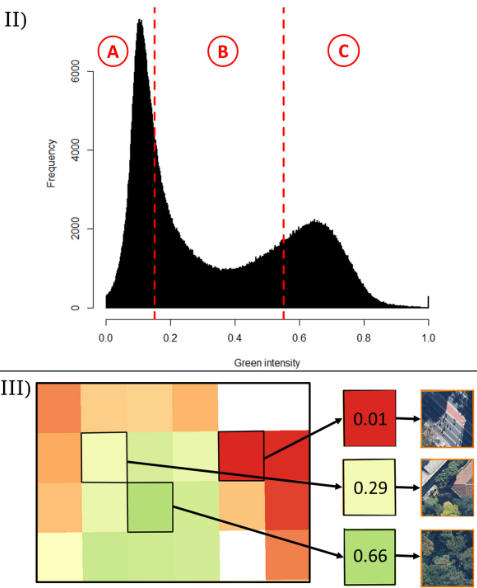
<https://essd.copernicus.org/articles/12/3545/2020/#section3>

### 

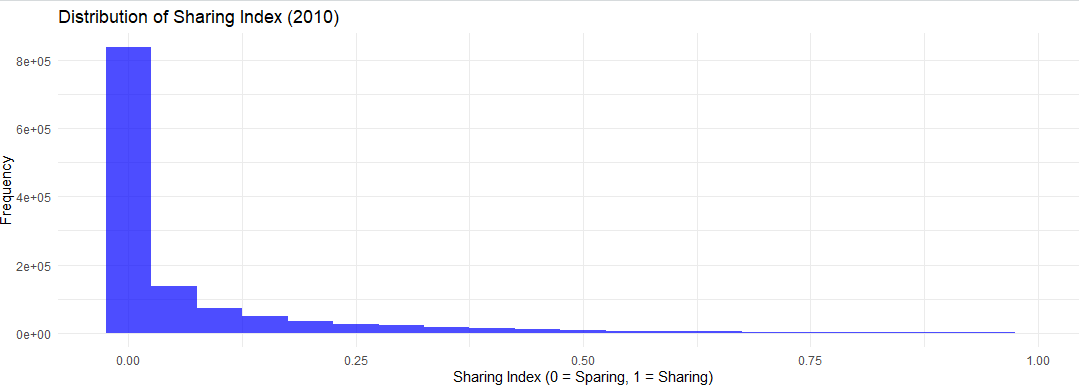
Anything that is not forest (tree cover)/cropland, switch to cultivated

Cropland + cultivated as argraicultural land



****

**Sharing\_index\_2010**

****

sum(values > 0.1) / length(values) \* 100

[1] 19.00413

81% sharing\_index\_2010 <0.1

**Agriculture change (proportion of pure crop + mixed crop change)**

**cat(sprintf("Proportion with small change (±0.1): %.2f%%\n", prop\_small\_change))**

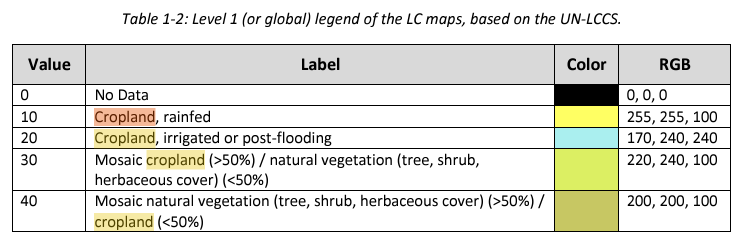
**Proportion with small change (±0.1): 98.60%**

**ESA Land cover change** <https://cds.climate.copernicus.eu/datasets>

* Land cover change data from 2005-2022 downloaded inY:\sparing\_sharing\Data\ESA\_LC
* Land cover class codes for cropland

Product user guide Page 14

<https://dast.copernicus-climate.eu/documents/satellite-land-cover/WP2-FDDP-LC-2021-2022-SENTINEL3-300m-v2.1.1_PUGS_v1.1_final.pdf>



IPCC Classes considered for the change detection (Agriculture *page 18 in the guide*)

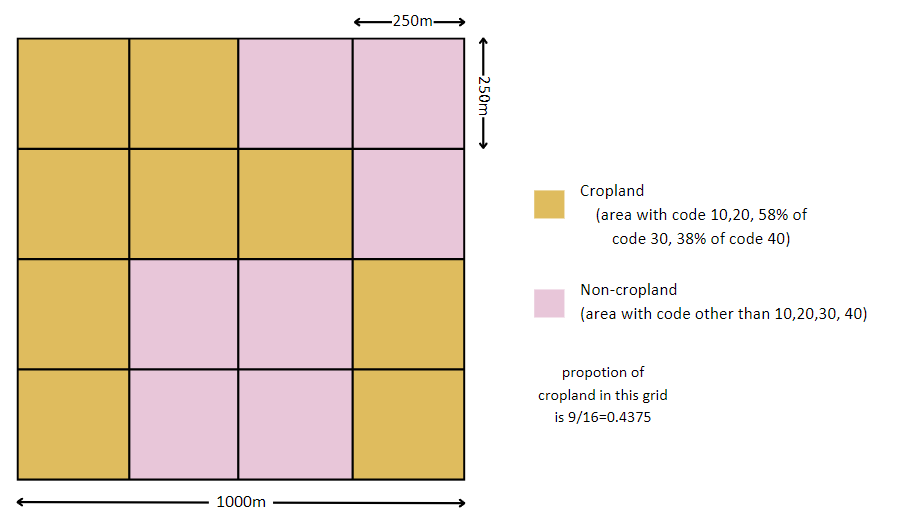
* Reclassify the raster only to retain those representing cropland (10, 20, 30, 40?), assigning all other classes a value of 0 (non-cropland). 300m cell

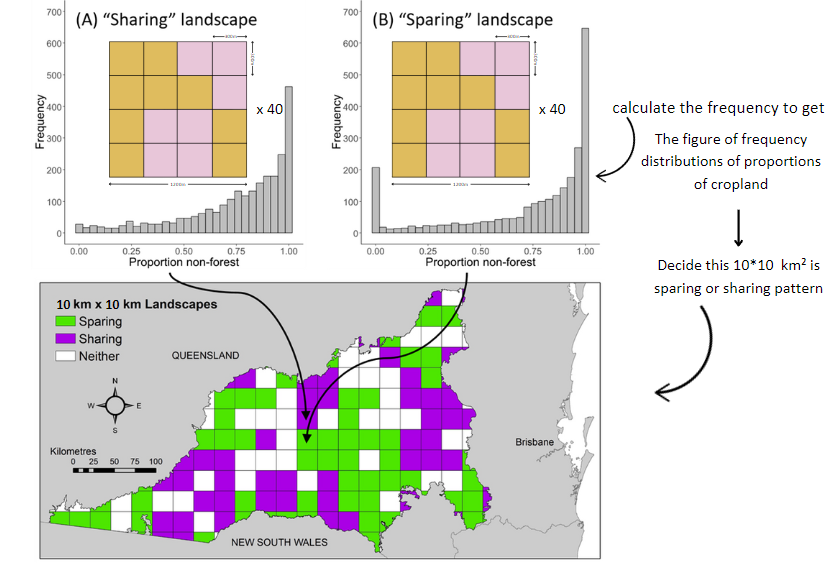
Land sparing landscapes will tend to have greater variation in land use intensities (due to a higher frequency of high or low land use intensities);

Land sharing will tend to have agricultural land use intensities that are roughly uniformly distributed.

* 58% of cropland in code 30 and 38% of cropland in code 40 pixels -The weights for GlobCover are 0.58 and 0.38, respectively, while the cropland cover in the mosaic cropland/vegetation type is 50–70% and in mosaic vegetation/cropland is 20–50%

<https://www.tandfonline.com/doi/full/10.1080/17538947.2013.854414#d1e942>





* Cropland intensification is characterised by the conversion of a mosaic of cropland and natural vegetation (class values 30 and 40) to a rainfed or irrigated cropland (class values 10 to 20) (*page 29 in the guide*)

**Spatial Production Allocation Model (SPAM 2020 V1r0)** <https://mapspam.info/>

Methodology paper <https://www.sciencedirect.com/science/article/pii/S0308521X14000110?via%3Dihub>

A cultivated planet in 2010 – Part 2: The global gridded agricultural-production maps:

<https://essd.copernicus.org/articles/12/3545/2020/#section3>

* Uses actual national or subnational yield data to downscale through a spatial allocation model at a finer resolution (10 km x 10 km)
* Uses an algorithm that estimates where specific crops are likely to be grown within each country based on suitability factors (e.g., soil, climate, market access, etc.)
* Yield data is available in 2000/2005/ 2010/2017/2020, for 46 crops including maize, rice, wheat, soybean <https://cgiardata.projects.earthengine.app/view/spam>
* Yield is a measure of productivity, the amount of production per harvested area, and is measured in **kilogram/hectare**. The total yield of a crop, when considering all production systems, is not the sum of the individual yields, but the weighted average of the 4 yields.
* Yield data downloaded under Y:\sparing\_sharing\Data\SPAM\_Yield

##### **Production (P)**

Production, for each production system and crop, is calculated by multiplying area harvested with yield (metric tons)

##### **Physical area (A)**

Physical area is measured in a hectare and represents the actual area where a crop is grown, not counting how often production was harvested from it

##### **Harvested area (H)**

Also measured in a hectare, harvested area is at least as large as physical area, but sometimes more, since it also accounts for multiple harvests of a crop on the same plot. Like for physical area, the harvested area is calculated for each production system and the sum of all harvested areas of all production systems in a pixel amount to the total harvested area of the pixel.

The sum of all the harvested areas of the crops in a pixel can be larger than the pixel size.

**Next Steps:**

So, resample the land use data to 250 m and snap to the yield data. Then, for each 10 km landscape represented by the yield data calculate the sparing/sharing pattern (maybe based on 4 x 4 blocks of 250 m land use grids) and amount of cropland.

Land cover data pre-processing:

1. Generate binary cropland-noncropland map in 2000, 2005, 2010 and 2020 (finished)

2. Convert cropland maps to a finer 250m resolution (finished)

3. Physical area data download 2000, 2005, 2010, 2020 (finished)

4. Harvested area data download 2000, 2005, 2010, 2020 (finished)

5. Create for each year - stack of physical area/ stack of harvested area / stack of yield for each crop (finished)

6. Sum up harvested area for each year

7. Sum up physical area for each year

8. Take mean for aggrefating yield area for each year (finished)

9. Find correct code for esa dataset, determine which codes are natural vegetation (finished)

10. Use esa datasets, 10/20 are high intensity agriculture, 30/40 mix areas are medium intensity, other natural vegetation are low intensity, get the frequency of cropland - sparing\_sharing index. (finished)

<https://besjournals.onlinelibrary.wiley.com/doi/full/10.1111/1365-2664.14195>

11. Make sharing\_index vs yield maps, covariate maps

Yield data pre-processing:

1. Create a single yield map combining all crops

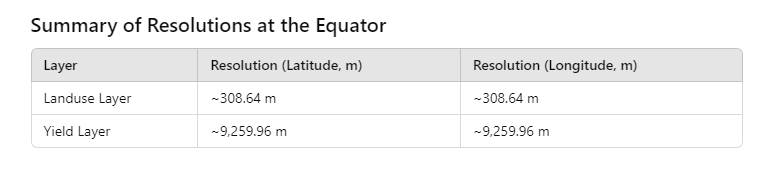
* Stack rasters to load all 42 layers into a single object
* Aggregate by sum or average?
* SPAM2000 - choose all TA (only 21 crops!)
* SPAM2005 - choose all TA (42 crops)
* SPAM2010 - choose all \_A (42 crops)
* SPAM2017 - not global, only Africa South of the Sahara! (42 crops) skip this one
* SPAM2020 - choose all \_A (46 crops?)

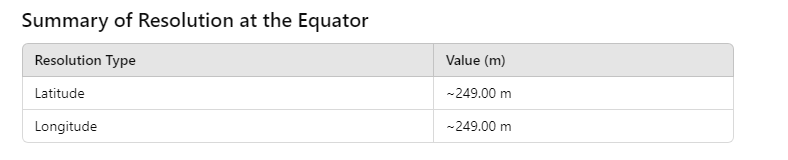
1. Snap/Align the resampled land use data with the yield data grid, so each cell corresponds spatially to the same area in both datasets.

Calculating:

**Based on 4 x 4 Blocks of 250m Cells**:

* Each **10km x 10km block** contains **1600 cells** (since 10km = 40 cells at 250m resolution).
* Divide the block into **4 x 4 sub-blocks** (each sub-block contains 16 cells).
* For each sub-block, calculate:
  1. **Proportion of Cropland**: Count the number of cropland cells (land use class for cropland) and divide by 16 (total cells in the sub-block).
  2. **Variance in Cropland**: Measure variability in cropland coverage between sub-blocks to determine the sparing/sharing pattern:
     + **High variance**: Indicates a land-sparing pattern (cropland is concentrated in some sub-blocks).
     + **Low variance**: Indicates a land-sharing pattern (cropland is spread evenly across sub-blocks).
     + Integrate yield data here?

Original yield and landuse layers: 

Resample binary cropland landuse layer:

The land cover map has a resolution of 300m x 300m. First, I will generate a 4x4 grid (forming an area of 1000m x 1000m) and calculate the proportion of cropland in each grid by dividing the number of cropland cells by the total number of cells.

I will then scale this up to a 10km x 10km landscape, which consists of around 40 grids, and calculate the frequency distribution of cropland proportions within this larger area.

Based on the frequency distribution, I will determine whether the 10km x 10km landscape follows a land sparing or land sharing pattern.

This process will be applied to the entire global map. (the following is a very rough workflow illustration, is that make any sense to you?

**Questions:**

##### #### The following are previous notes #####

**Datasets:**

1. Land cover change datasets:

Global cropland 3km expansion map (2000-2019) [(Potapov et al. 2022)](https://paperpile.com/c/oIhvJ6/ugxD)

<https://glad.umd.edu/dataset/croplands>

* From 2003 to 2019, cropland area increased by 9% and cropland net primary production (NPP) by 25%, primarily due to agricultural expansion in Africa and South America.
* Half of the new cropland area (49%) replaced natural vegetation and tree cover, indicating a conflict with the sustainability goal of protecting terrestrial ecosystems.
* From 2003 to 2019, global per-capita cropland area decreased by 10% due to population growth. However, the per-capita annual cropland NPP increased by 3.5% as a result of intensified agricultural land use.
* 3km resolution too coarse

1. Yield datasets:

The global gridded agricultural-production maps [(Yu et al. 2020)](https://paperpile.com/c/oIhvJ6/lNbo) <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/PRFF8V>

* Model based on the 2009–2011 average of the crop production statistics (no dynamic changes).

Production/Yield quantities of crops and livestock products in World 1994 - 2022 [(FAOSTAT 2024)](https://paperpile.com/c/oIhvJ6/Ttzz)

<https://www.fao.org/faostat/en/#data/QCL/visualize>

* Individual crop/livestock product yield data available from 1994-2022 (allowing for analysis of dynamic changes over time)
* High-demand products like cereals or cash crops (e.g., palm oil, soy, rice, wheat, and maize - major cereal crops)
* Only country-scale data

EarthStat (<http://www.earthstat.org/>)

* Latest until 2008 (out of date)
* Yield Trends and Changes for Maize, Soybean, Rice, and Wheat (could be a reference <http://www.earthstat.org/yield-trends-changes-maize-soybean-rice-wheat/> )

Global Yield Gap Atlas (<https://www.yieldgap.org/>)

* Only country-level data (too coarse)

FAO's Global Agro-Ecological Zones (GAEZ) (<http://www.earthstat.org/>)

* Country totals are based on FAO statistics for the years 2009-2011
* <https://gaez.fao.org/pages/theme-details-theme-5> (Actual Yields and Production)

EarthStat (<http://www.earthstat.org/>)

* Latest until 2008 (out of date)
* Yield Trends and Changes for Maize, Soybean, Rice, and Wheat (could be a reference <http://www.earthstat.org/yield-trends-changes-maize-soybean-rice-wheat/> )

**Spatial Production Allocation Model (SPAM 2020 V1r0)** <https://mapspam.info/>

* Uses actual national or subnational yield data to downscale through a spatial allocation model at a finer resolution (10 km x 10 km)
* Uses an algorithm that estimates where specific crops are likely to be grown within each country based on suitability factors (e.g., soil, climate, market access, etc.)
* Yield data is available in 2005/ 2010/2017/2020, for 46 crops including maize, rice, wheat, soybean <https://cgiardata.projects.earthengine.app/view/spam>

Other proxy could be:

Distance to markets, processing facilities, and cities;

Biophysical variables (i.e., fraction of the Photosynthetically Active Radiation <https://link.springer.com/article/10.1007/s11119-022-09970-8>) soil/climatic

**Papers**:

The global cropland-sparing potential of high-yield farming [(Folberth et al. 2020)](https://paperpile.com/c/oIhvJ6/sQPZ)

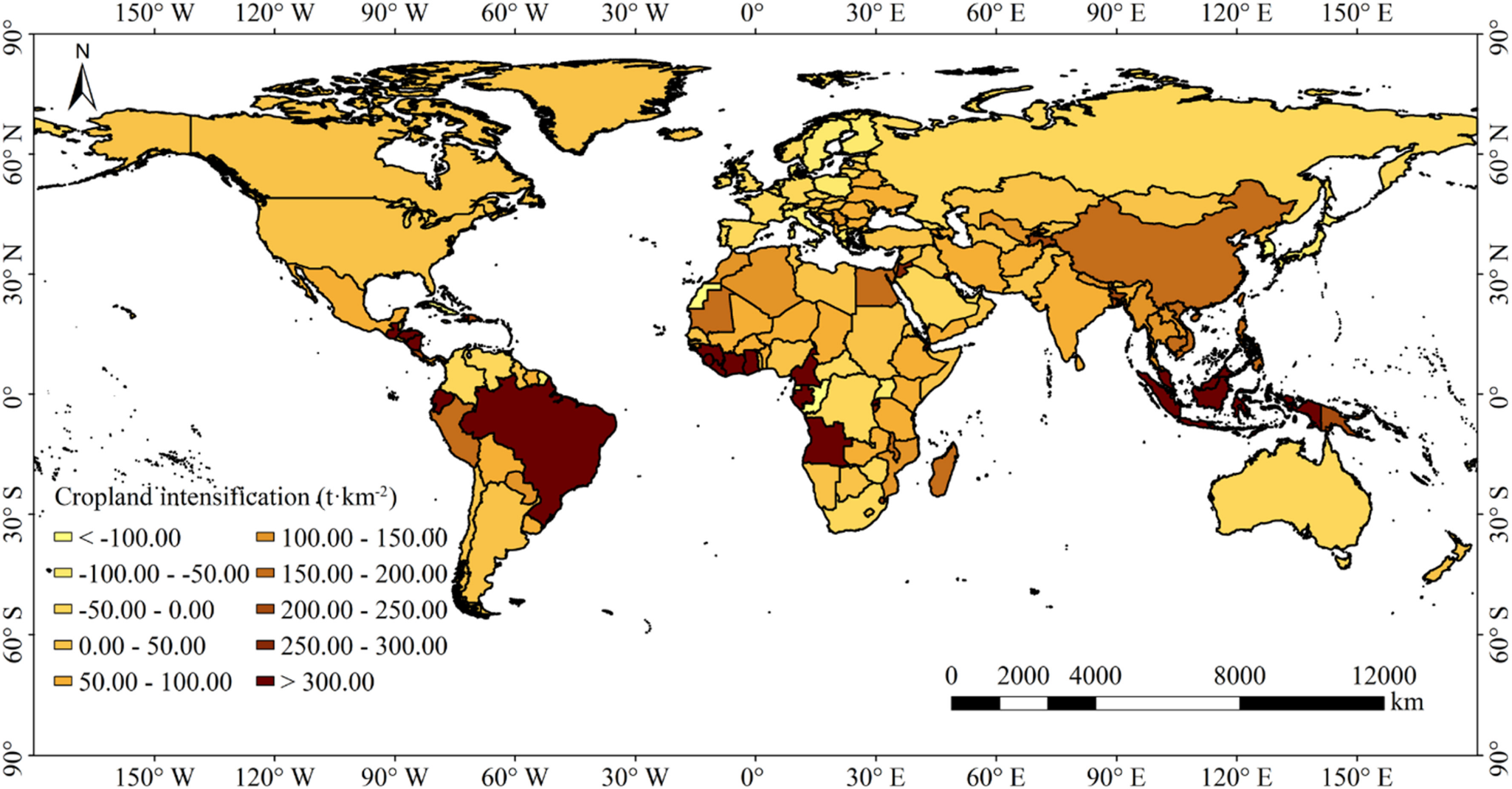
* Crop yields from the EPIC-IIASA global gridded crop model (or statistically derived yield datasets (FAOSTAT))
* Cropland data from SPAM 2005 (or other suitable land-use datasets)
* Global Spatially-Disaggregated Crop Production Statistics Data for 2005 Version 3.2 (IFPRI and IIASA, 2016)
* Downloaded supplementary data located at Y:\sparing\_sharing\Data\Folberth\_2020\_Supplementary\_Data\_1

A New Indicator for Global Food Security Assessment: Harvested Area Rather Than Cropland Area [(Song et al. 2022)](https://paperpile.com/c/oIhvJ6/QEkk)

* Global land use and land cover (LULC) data from 1992 to 2015 from the European Space Agency (ESA), with a spatial resolution of 300 m
* Global grain yield and the grain yields of 30 countries with large cropland areas from 1992 to 2015 from the Food and Agriculture Organization Corporate Statistical Database (FAOSTAT)
* Paper downloaded at Y:\sparing\_sharing\Papers

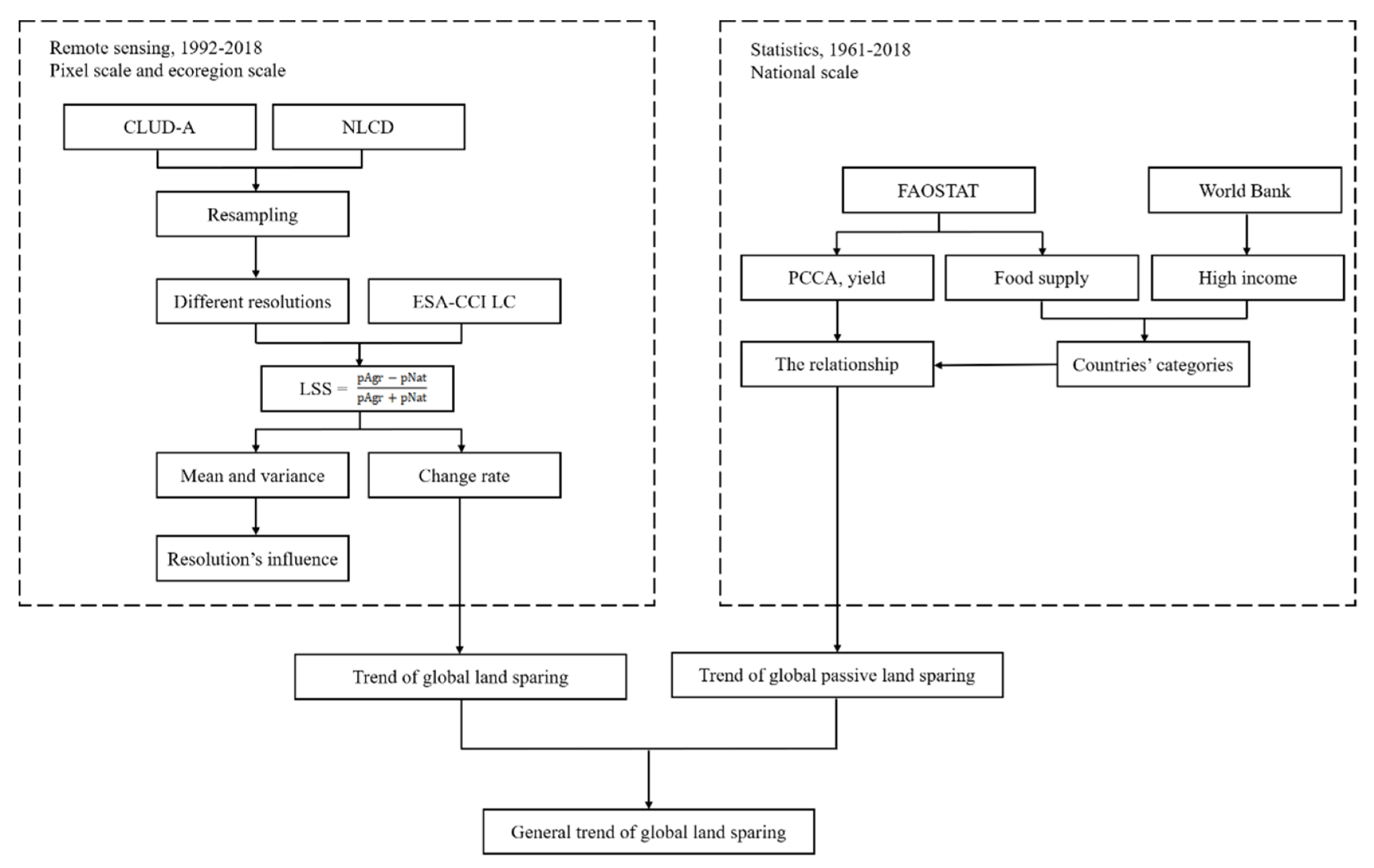
Global cropland intensification surpassed expansion between 2000 and 2010: A spatio-temporal analysis based on GlobeLand30 [(Hu et al. 2020)](https://paperpile.com/c/oIhvJ6/34R2)

* Found that Brazil ranked first in cropland intensification, followed by China, India and Ukraine.



Global Change of Land-Sparing and Land-Sharing Patterns over the Past 30 Years: Evidence from Remote Sensing and Statistics [(Zhao et al. 2021)](https://paperpile.com/c/oIhvJ6/kf0Z) <https://www.mdpi.com/2072-4292/13/24/5090>



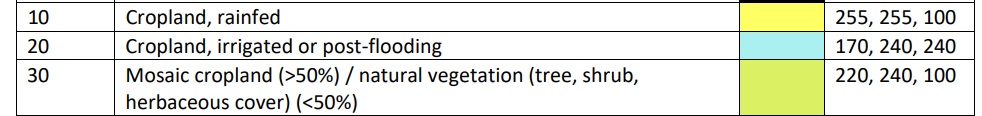
* The index compares the balance between agricultural land and natural habitats. A positive LSS indicates more land is used for agriculture (land sparing pattern), while a negative LSS suggests a higher proportion of natural habitat (land sharing pattern)
* The calculation of the LSS index is based on the global ESA-CCI LC, which is a time series of consistent global LC maps, with a 300-m spatial resolution, from 1992 to 2018
* The actual integration of yield data happens when they analyse agricultural intensification trends globally, but it is not part of the LSS formula itself

Balanced spatial distribution of green areas creates healthier urban landscapes [(Cirino et al. 2022)](https://paperpile.com/c/oIhvJ6/lUEj) <https://www.researchgate.net/publication/360458205_Balanced_spatial_distribution_of_green_areas_creates_healthier_urban_landscapes>

* Divides the landscape into quantiles of green cover and urbanization using the Enhanced Vegetation Index (EVI) from LANDSAT data

ESA <https://cds.climate.copernicus.eu/datasets>

Product user guide Page 14 <https://datastore.copernicus-climate.eu/documents/satellite-land-cover/WP2-FDDP-LC-2021-2022-SENTINEL3-300m-v2.1.1_PUGS_v1.1_final.pdf>



reclassify the raster to only retain those representing cropland, assigning all other classes a value of 0 (non-cropland). This will create a binary cropland/non-cropland map.

// LULC classification

The latest global LULC classification divides surface

objects into seven categories: cropland, forestland,

grassland, wetland, urban construction land, unused land

and water bodies. Cropland refers to land where crops

are planted, including cropland, newly developed, reclaimed and arranged land, and leisure land (such as rotation land and rotation land). The land is planted

mainly with crops (including vegetables), with sporadic

fruit trees, mulberry trees, or other trees. Cultivated

beaches and sea beaches that can ensure a harvest for

one season on average every year also have been considered to be cropland (Portmann et al., 2010). [Song et al. 2022)](https://paperpile.com/c/oIhvJ6/QEkk)//

Next steps:

1. Download ESA land use data, standardise cropland categories, and exclude urban areas and other non-relevant land types
2. Calculate the proportion of cropland versus natural habitat within each 300m cell using ESA data
3. Plot the frequency of the proportion non-forest
4. Download SPAM yield data for the years 2005, 2010, 2017, and 2020.
5. Papers to look at: <https://onlinelibrary.wiley.com/doi/full/10.1111/j.1365-2486.2009.01849.x?casa_token=vnkhvvSM4TwAAAAA%3Ad0LWn0dWDWxZMBV5smqhHveLDpWKSFgYtIbIahF7vv47049SPFRcXxmPrfe8Vk2flrYVF-FKsoR7598K>

<https://www.nature.com/articles/s41893-023-01073-0>